CAN BIOMECHANICS CONTRIBUTE TO CLINICAL ORTHOPAEDIC ASSESSMENTS?

Richard A. Brand, M.D.

Department of Orthopaedic Surgery
The University of Iowa
Iowa City, Iowa 52242

Keynote Address
International Society of Biomechanics
Amsterdam
July, 1987

I wish to ask the question, "Can biomechanics contribute to clinical orthopaedic assessments?" This question is important because many articles answer this question affirmatively, even emphatically without compelling evidence or data to support their contention (Strong opinions without data often sway the bulk of medical opinion). It is fair to say that my opinions regarding clinical "gait analysis" are with the minority among locomotion researchers and may even be controversial. However, I believe the bulk of evidence is in my favor.

It is important to state my philosophical notions about medicine, because those notions are inextricably linked with my ideas about medical "tests," regardless of type. Furthermore, the constructive intent might be misinterpreted without grasping its underlying philosophy. My approach might be called an epistemological one, since the process is one of identifying the methods used to gain useful clinical knowledge.

Three philosophical notions affect my interpretation of diagnostic test use. (Aristotle supposedly said, "philosophy begins in medicine, and medicine ends in philosophy.") First, healers and practitioners of the medical arts have for all of recorded history, at least in the Western world, practiced what I would call "interventionism". The perception of "abnormality" demands intervention to restore "normality". Whether such expectation arose more from society or healers is not clear, but both contribute significantly. For practitioners, intervention is a livelihood. Also, the burden of proving efficacy lies with the practitioner who chooses not to intervene, rather than with the practitioner who intervenes; a situation I find logically backward, and ethically despicable. I mention interventionism because I view use of diagnostic and assessment tools as a form of interventionism, and because I believe the researcher who claims clinical applicability of some tool has the responsibility to document that his or her measures provide truly useful and unique, not simply interesting information. Just as a treatment should be proven efficacious, so should a new test.

Second, few would dispute that medical care costs are rising in all industrialized countries. Perhaps the most egregious example is in the United States, where approximately eleven percent of our enormous gross national product is expended on what is inappropriately called "health care". I feel this is more appropriately termed "illness intervention" which is, at best, indirectly related to the health of a society. No end is in sight for this rise in cost, and I see little likelihood that we will direct our resources to health, rather than to illness. A single factor responsible for such a complex and paradoxical phenomenon cannot be identified, but no one could question the role of technological "advances" and our indiscriminate worship of technology. Given the cost of "illness care", any new medical test must be unique, useful and demonstrably cost-effective.

Third, "biomechanics" has emphasized "mechanics" and has largely viewed its biologic subjects as structures with complex geometry and static physical properties, rather than living organisms. Fluxes in geometry, properties and behavior are usually ignored. However, such fluxes ignored by the clinician result in a most unsatisfactory patient-physician encounter. The more useful tests will account for and be conceptually consistent with the dynamic nature of the human organism.

With that background, let me return to biomechanics and clinical tests. I will focus exclusively on the role that biomechanics plays with patients. I am not considering how biomechanics can contribute to our concepts of disease; we all recognize that biomechanics can answer well-formulated questions, sometimes using groups of patients for appropriate observations. I am most familiar with studies of locomotion and will focus upon these. However, I do not intend to exclude other mechanical evaluation methods, and will mention a few examples. Regardless of my focus, my remarks are applicable to other mechanical measures in addition to gait analysis.

A clinician orders a "test" on a patient for one of four reasons: to distinguish between disease entities (i.e., diagnosis); to determine the severity of disease or injury (i.e.,

assessment or evaluation); to select among several treatment options; and to predict prognosis. There are no other patient-related reasons for ordering a clinical test.

Biomechanical measures of locomotor function were introduced over one hundred years ago; however, none have gained widespread clinical use for individual patients. Every day, clinicians routinely order technologically sophisticated tests such as hemoglobin levels, x-rays, serum glucose levels or magnetic resonance imaging, but no "biomechanical" tests are routinely used. Biomechanical measures have largely been technology looking for applications. They remain parochial tools awaiting proof of clinical usefulness, regardless of cleverness or technological sophistication; sophistication is no assurance of usefulness. The establishment of criteria for selecting and developing potential biomechanical measures and the exploration of strategies validating these measures enhance their potential clinical value.

Upon critical reflection, most reported biomechanical measures do not adequately contribute to the four reasons for ordering tests. Many explanations may be cited: most measures can be observed by an astute clinician; many mechanical measures are dependent upon the patient's mood, motivation, performance or pain; some measures exhibit short-term variability rather than a desirable stability; and few have been independently validated.

My former colleague, Roy Crowninshield, and I were thinking about this issue several years ago and came up with several usefulness criteria for biomechanical measures (Brand and Crowninshield, 1981). Since that time, I have added several more criteria to the list. At the present, I think the following constitutes reasonable criteria when developing or selecting a biomechanical measure: it must be accurate and reproducible; the measurement technique must not significantly alter the function it is measuring; it should exhibit reasonable stability; the measure should not be directly observable by the skilled clinician; the measure should be independent of mood, motivation or pain; it must clearly distinguish between normal and abnormal; the measure should be reported in a form analogous to some accepted clinical concept; the measure should be cost-effective; and finally, it must be appropriately validated. I do not assume that my list is all inclusive; others might propose equally or more important criteria.

Let me amplify several of these points. Few would argue with the first two criteria. To me, the most useful tests are those providing measures I can not observe. Being able to quantify what can be determined by simple observation adds very little, although there are a few exceptions. A patient's blood glucose, carbon dioxide content of arterial blood and the presence of a bone cyst represent findings which I could only suspect with my powers of observation. So it is with biomechanical measures; the most useful biomechanical measures will be those providing measures we can not observe, yet the emphasis has been quite the contrary. I hasten to reiterate that such quantitative distinctions assume far more importance in large groups of patients when answering some clearly posed question than in evaluating a single patient. Measures are likely to be more helpful when they are not dependent upon the patient's mood, motivation or level of pain. Tests such as blood glucose meet this criteria; measures such as temporal and distance factors of gait do not.

Measures are also more valuable when there is clarity between normal and abnormal. When there is considerable overlap of normal and abnormal, such as in temporal and distance factors of gait, the measure has far less value and can be interpreted only with corollary information. I will cite two examples of this phenomenon. When we were doing a long-term follow-up study of patients with clubfoot, it seemed logical to "evaluate" their gait (Brand, Laaveg, Crowninshield, & Ponseti, 1981). This we did in several conventional ways, but we additionally determined foot-floor pressures and the location of the center-of-pressure path. There were virtually no surprises. Pressure abnormalities could have been determined far more simply by looking at calluses on the patient's feet. Center-of-pressure paths for these clubfoot patients, while more variable than normals, usually fell within a range of normal. The center of pressure paths also failed to distinguish between dysfunctional feet and feet with only radiographic deformity. Schneider and Chao (1983) reached similar conclusions when looking at patterns of foot-floor reactions in patients with total knee replacements. Apparently, there was little difference between the reactions of normal subjects and patients with total knee replacements. However, when they performed a Fourier analysis they found distinctions between normal and abnormal. This suggests that sometimes the problem is not the measure but the analysis.

A measure is more useful when described in clinically relevant terms and concepts. Any measure requiring new concepts will be difficult to introduce and will likely meet with failure.

Cost-effectiveness does not directly relate to medical efficacy, and can not be the cause for clinical failure of biomechanical measures. However, cost-effectiveness will become increasingly more important in the application of any new measures. The measuring tool should reduce costs by eliminating unnecessary treatment, or by identifying conditions early and avoiding expensive complications.

Validation is critical in establishing medical efficacy. Valid measures may predict a different outcome than would have been previously predicted or suggest a different treatment than would have been recommended. If a measure does not change our predictions about a disease course or it does not change our recommendations for treatment, it is not valid (i.e., useful).

After reviewing the literature, it seems obvious that few measures are actually validated. Some years ago, Roy Crowninshield and I went through our files to classify papers on gait analysis. Our classification scheme was certainly not rigorously scientific, nor was our sample of papers necessarily random; nonetheless, subsequent reading has not changed my opinion about our conclusions: most papers report fascinating technology, but few document clinical usefulness (Table 1). Validation is so critical that any new measure should be designed with this in mind. A measure which a *priori* can not be validated is doomed to failure.

Table 1
GAIT PAPERS
FILES OF RAB AND RDC

Classifications (major purpose)	Number of Papers	Percent of Total
Description of Experimental or Numerical Method	65	44
Descriptive Results	48	32
Proposed Application	12	8
Demonstrated Application	8	6
Validated Application	0	0
	146	100

In the previously mentioned clubfoot study we attempted validation of the biomechanical measure (i.e., location of the center of pressure path) against an independent measure, a functional rating scale conceptually similar to many hip and knee rating systems (Brand et al., 1981). However, our biomechanical measure correlated neither with the functional rating scale nor radiographic changes. Thus, the biochemical measure failed to distinguish normal from abnormal, did not demonstrate anything I could not already observe, and it could not be validated against independent measures.

In a more successful effort, Chao and his colleagues reported what they termed a "performance index", based on the gait patterns of patients with knee disability (Chao, Laughman and Stauffer, 1980). They used stepwise discriminate analysis to select and weigh seven gait variables out of a potential 43 to create their performance index. This procedure usually discriminated between normal and abnormal. Their approach was promising, but contained several unsolved problems. First, the performance index correlated with a Harris Hip Score when all normals and abnormals were considered. However, the performance index did not discriminate knee disability among abnormal patients. Second, in another study of 254 patients considering 54 candidate variables, nine, rather than seven variables, were selected and their relative weighting changed (Laughman, Stauffer, Ilstrup and Chao, 1984). This means that weighting is sensitive to the data base. If such measures are so sensitive, we either need measures less sensitive to the data base or much larger data bases. Third, the weighting coefficients were different for men and women, again demonstrating a different sensitivity of subgroups. Recently, when Chao and colleagues developed a performance index for hip disability, the weighting factors again changed, showing sensitivity to selection of joint (Kaufman, Chao, Cahalan, Askew and Bleimeyer, 1987). This approach holds considerable promise, but it would ideally be less sensitive to such variables. It also requires more extensive validation against independent measures, such as clinical scores.

The minimal elements of appropriate validation include: well-designed cross-sectional and longitudinal clinical trials with statistically significant numbers of normals and abnormal patients, use of appropriate statistical methods to sort out potentially confounding patient variables, correlation of the measure with accepted independent measures, validation studies from several independent institutions.

Cross-sectional and longitudinal trials accomplish two things. First, they tell us whether the measure distinguishes normal from abnormal, or treated from untreated patients. Second, they tell us whether the measure fulfills one of the four criteria of usefulness mentioned earlier. Properly designed, prospective cross-sectional and longitudinal trials are the only way usefulness can be documented. Such trials require adequate normative data controlling for age, height, weight, time since injury, and so forth; hundreds or thousands of normals may be needed. Cross-sectional studies are valuable because of their ease, but are flawed by their poor control. They are most suited for the early phase of validation. Longitudinal studies are more valuable for validation, because each patient serves as his or her own control. In longitudinal studies, changes in a measure reflect variation in a disease process due to treatment or time. Only with longitudinal trials can we tell whether a measure suggest one treatment option over another. However, longitudinal studies are difficult and expensive to perform. These problems with longitudinal studies do not diminish the responsibilities incurred by those who introduce new measures. Good examples of appropriate longitudinal studies are those by Jernberger (1970) and more recently by Kenwright et al. (1986), who used stiffness as a measure of fracture healing.

Biomechanical investigators have recently developed measures with more sophisticated statistical tools. Grouping and weighting a number of measures may prove far more valuable than using single measures. Wong, Simon, and Olshen (1983) and Wooten, Kadaba, and Cochran (1984) used cluster analysis to more clearly distinguish normal from abnormal. Yamamoto and his colleagues (1983) reported principle component analysis of ten candidate gait variables to better distinguish normals from abnormals.

Can biomechanics contribute to clinical orthopaedic assessments? Currently, I must answer with a qualified yes. An example is seen in cerebral palsy where the degree of apparent weakness of the peroneal muscles does not necessarily correlate with the degree of peroneal muscle dysfunction. Nor does the presence of a strong muscle group on manual testing insure that the same muscle group will be normally active during gait. Owing to this lack of correlation, the results of tendon transfer surgery are unpredictable. Performing EMG's during gait and during other activities demonstrates which muscles are active and which are inactive, thus elucidating something we can not observe (Hoffer and Perry, 1983). Current published studies must still be considered preliminary; more definitive longitudinal validation studies are yet to be published.

Another example is a biomechanical test preoperatively predicting the outcome of a surgical procedure, such as high tibial osteotomy, an unpredictable and therefore seldom performed operation. Prodromos, Andriacchi, and Galante (1985) demonstrated that patients with unicompartmental knee arthrosis could be classified into two groups: those with a low adduction moment, and those with a high adduction moment. In a medium-term follow-up study (average 3.2 years), they demonstrated that the group with a low preoperative adduction moment did significantly better than those with a high adduction moment. Unfortunately, longer follow-up is lacking, as is independent verification from another laboratory. Nonetheless, this study is an excellent example of a biomechanical test contributing to clinical orthopedic assessments.

In conclusion, not all biomechanical measures will meet each of the usefulness criteria, nor are all of the criteria equally important. Also, others might formulate equally or more important criteria. Biomechanical measures intended to be clinically applicable must be developed with an appropriate validation strategy. In the future, adequate criteria for selection and validation of biomechanical measures can be met, and these measures can become clinically useful.

BIBLIOGRAPHY

1. Brand, R.A., and Crowninshield, R.D.: Comment on Criteria for Patient Evaluation Tools. J. Biomech., 14:655, 1981.

2. Brand, R.A.; Laaveg, S.J.; Crowninshield, R.D.; and Ponseti, I.V.: The Center of Pressure Path in Treated Clubfeet. Clin. Orthop. Rel. Res., 160:43-47, 1981.

3. Chao, E.Y.; Laughman, R.K.; and Stauffer, R.N.: Biomechanical Gait Evaluation of Pre and Postoperative Total Knee Replacement Patients. Archives of Orthopaedic and Traumatic Surgery, 97:309-317, 1980.

4. Hoffer, M.M., and Perry, J.: Pathodynamics of Gait Alterations in Cerebral Palsy and the Significance of Kinetic Electromyography in Evaluating Foot and Ankle Problems. Foot and Ankle, 4:128-134, 1983.

5. Jernberger, A.: Measurement of Stability of Tibial Fractures. A Mechanical Method. Acta Orthop. Scand., Suppl. 135, 1970.

6. Kaufmann, K.R.; Chao, E.Y.S.; Cahalan, T.D.; Askew, L.J.; and Bleimeyer, R.R.: Development of a Functional Performance Index for Quantitative Gait Analysis. In Biomedical Sciences Instrumentation. Edited by J.D. Enderle, pp. 49-55. Research Triangle Park, North Carolina: Instrument Society of America. 1987.

7. Kenwright, J.; Richardson; J.B., Goodship, A.E.; Evans, M.; Kelly, D.J.; Spriggins, A.J.; Newman, J.G.; Burrough, S.J.; Harris, J.D.; and Rowley, D.I.: Effect of Controlled Axial Micromovement on Healing of Tibial Fractures. Lancet, 2:1185-1187, 1986.

8. Laughman, R.K.; Stauffer, R.N.; Ilstrup, D.M.; and Chao, E.Y.S.: Functional Evaluation of Total Knee Replacement. J. Orthop. Res., 2:307-313, 1985.

9. Prodromos, C.C.; Andriacchi, T.P.; and Galante, J.O.: A Relationship Between Gait and Clinical Changes Following High Tibial Osteotomy. J. Bone and Joint Surg., 67-A:1188-1193, 1985.

10. Schneider, E., and Chao, E.Y.: Fourier Analysis of Ground Reaction Forces in Normals and Patients with Knee Joint Disease. J. Biomech., 16:591-601, 1983.

11. Wong, M.A.; Simon, S.; and Olshen, R.A.: Statistical Analysis of Gait Patterns of Persons with Cerebral Palsy. Statistics in Medicine, 2:345-354, 1983.

12. Wooten, M.E.; Kadaba, M.P.; and Cochran, G.V.B.: Feature Extraction and Prototype Selection for Normal Electromyographic Data in Gait Analysis. Transactions of the Orthopaedic Research Society, 9:158, 1984.

13. Yamamoto, S.; Suto, Y; Kawamura, H.; Hashizume, T.; and Kakurai, S.: Quantitative Gait Evaluation of Hip Diseases Using Principal Component Analysis. J. Biomech., 16:717-726, 1983.